The Annals of Probability
1987, Vol. 15, No. 1, 281–291

EXTREME VALUES FOR STATIONARY AND MARKOV SEQUENCES[1]

BY GEORGE L. O'BRIEN

York University

Let $(X_n)_{n=1,2,\dots}$ be a strictly stationary sequence of real-valued random variables. Let $M_{i,j} = \max(X_{i+1}, \dots, X_j)$ and let $M_n = M_{0,n}$. Let (c_n) be a sequence of real numbers. It is shown under general circumstances that $P[M_n \le c_n] - (P[X_1 \le c_n])^{nP[M_{1,p_n} \le c_n | X_1 > c_n]} \to 0$, for any sequence (p_n) satisfying certain growth-rate conditions. Under suitable mixing conditions, there exists a distribution function G such that $P[M_n \le c_n] - (G(c_n))^n \to 0$ for all sequences (c_n). These theorems hold in particular if (X_n) is a function of a positive Harris Markov sequence. Some examples are included.

1. Introduction. Let $X = (X_n)_{n=1}^{\infty}$ be a strictly stationary real-valued stochastic sequence. Let F be the distribution function of X_1 and let $x_F = \sup\{x\colon F(x) < 1\}$. Let $M_{i,j} = \max(X_{i+1}, X_{i+2}, \dots, X_j)$ and let $M_i = M_{0,i} = \max(X_1, X_2, \dots, X_i)$. The purpose of this paper is to investigate the asymptotic behaviour of $P[M_n \le c_n]$ as $n \to \infty$, for real-valued sequences $(c_n)_{n=1}^{\infty}$. We always assume $P[X_1 > c_n] > 0$ and $P[X_1 = x_F] = 0$.

If X is a sequence of independent identically distributed random variables (i.i.d. sequence), then, of course,

$$P[M_n \le c_n] = (F(c_n))^n. \tag{1.1}$$

Our purpose may be thought of as an investigation of the asymptotic effect of the dependence structure of X on $P[M_n \le c_n]$. We will show in Theorem 2.1 that under a broad class of circumstances we have

$$P[M_n \le c_n] - (F(c_n))^{nP[M_{1,p_n} \le c_n | X_1 > c_n]} \to 0, \tag{1.2}$$

where $(p_n)_{n=1}^{\infty}$ is any sequence of positive integers satisfying $p_n = o(n)$ and certain other rate of growth conditions.

Comparing (1.1) and (1.2), we see that M_n is asymptotically no larger in the general case than in the i.i.d. case and the conditional probability in (1.2) provides a measure of how much smaller it tends to be. This conditional probability should be interpreted as a measure of the extent to which large values cluster together, since M_{1,p_n} can be less than c_n when $X_1 > c_n$ only if X_1 is the last element in a cluster of values which exceed c_n. If large values

Received July 1984; revised June 1986.
[1]This research was supported in part by the Natural Sciences and Engineering Research Council of Canada. It was mainly done while the author was being supported by the School of Operations Research and Industrial Engineering and the Center for Applied Mathematics at Cornell University.
AMS 1980 *subject classifications.* Primary 60F05, 60G10; secondary 60J05.
Key words and phrases. Extreme value, stationary sequence, maximum, minimum, weak limit, mixing, extremal index, phantom distribution function, function of a Markov sequence.

do cluster together, there must be longer intervals between clusters, so that $P[M_n \le c_n]$ is larger than it would be otherwise.

If $P[M_{1,p_n} > c_n | X_1 > c_n] \to 0$, then (1.2) reduces to

$$P[M_n \le c_n] - (F(c_n))^n \to 0.$$

This result has been proved under assorted conditions by Loynes (1965), Leadbetter (1974), O'Brien (1974a) and Davis (1979), among others.

If $\Gamma(m) := \limsup_{n\to\infty} P[M_{m,p_n} > c_n | X_1 > c_n] = 0$ for some fixed m, then (1.2) yields $P[M_n \le c_n] - (F(c_n))^{nP[M_{1,m} \le c_n | X_1 > c_n]} \to 0$, which has been proved under various conditions by Newell (1965), O'Brien (1974b) and Leadbetter (1983). The last two papers also deal with the case $\Gamma(m) \to 0$ as $m \to \infty$. All these results can be deduced from our Theorem 2.1. Rootzén (1986) has independently obtained a result much like ours.

The main theorem and some related results are given in Section 2 and proved in Section 3. In Section 4, we consider the simultaneous asymptotic behaviour of $P[M_n \le c_n]$ for collections of sequences (c_n).

Many examples from the literature on extreme values can be viewed as functions of Markov sequences. We give a systematic study of these sequences in Section 5.

The notation introduced in the first paragraph will be used throughout this paper. We also make the following notational conventions. Let (b_n) be a sequence. To avoid many subscripts, we often write b for b_n. Limits are understood to be "as $n \to \infty$." We write $F^k(\cdot)$ instead of $(F(\cdot))^k$ and $P^k[\cdot]$ for $(P[\cdot])^k$. If x is a real number, $[x]$ denotes the greatest integer not exceeding x.

2. The main theorem. We will use the following mixing condition, which weakens Leadbetter's (1974) widely used condition $D(c_n)$ just enough to cover the case of certain stationary sequences with a periodic structure, such as periodic Markov chains.

DEFINITION 2.1. The stationary sequence (X_n) is said to have asymptotic independence of maxima relative to the sequence (c_n) of real numbers [notation: (X_n) has AIM(c_n)] if there exists a sequence (q_n) of positive integers with $q_n = o(n)$ such that

$$\alpha_n := \max\left| P\left[M_i \le c_n,\, M_{i+q,i+q+j} \le c_n\right] - P\left[M_i \le c_n\right] P\left[M_j \le c_n\right] \right| \to 0, \tag{2.1}$$

where the maximum is over all $i \ge q$ and $j \ge q$ such that $i + q + j \le n$.

THEOREM 2.1. *Let (X_n) have* AIM(c_n), *where (c_n) is a sequence of real numbers. Let (q_n) and (α_n) be as in Definition* 2.1. *Let (p_n) be a sequence of positive integers such that*

$$p = o(n), \qquad n\alpha = o(p) \quad \textit{and} \quad q = o(p). \tag{2.2}$$

If either

$$\liminf F^n(c_n) > 0 \tag{2.3}$$

or

$$\liminf P\left[M_{1,p} \le c_n | X_1 > c_n\right] > 0, \tag{2.4}$$

then both

$$P[M_n \le c_n] - \left(F(c_n)\right)^{nP[M_{1,p} \le c_n | X_1 > c_n]} \to 0 \tag{2.5}$$

and

$$P[M_n \le c_n] - \exp\left\{-nP\left[X_1 > c_n, M_{1,p} \le c_n\right]\right\} \to 0. \tag{2.6}$$

The proof of this theorem also yields the following corollaries.

COROLLARY 2.2. *If the assumptions of the theorem other than* (2.3) *and* (2.4) *hold, then we still have*

$$P[M_n \le c_n] \le \exp\left\{-nP\left[X_1 > c_n, M_{1,p} \le c_n\right]\right\} + o(1). \tag{2.7}$$

COROLLARY 2.3. *Let* (k_n) *be an increasing sequence of positive integers. Then Theorem* 2.1 *and Corollary* 2.2 *hold if* n *is replaced by* k_n *throughout their statements.*

REMARKS. If (a_n) is any sequence of elements of $[0,1]$, then

$$a_n^n - \exp\{-n(1-a_n)\} \to 0. \tag{2.8}$$

This can be proved by applying Theorem 1.5.1 in Leadbetter, Lindgren and Rootzén (1983) to subsequences along which a_n^n converges. In particular, (2.3) is equivalent to the condition $\limsup nP[X_1 > c_n] < \infty$. Also, (2.5) and (2.6) are equivalent if either (2.3) or (2.4) holds. The requirement that (2.3) or (2.4) holds serves to preclude the delicate situation where the second term in (2.5) converges, so to speak, to 0^0. At the cost of a much messier proof, (2.6) can be proved with (2.3) replaced in the theorem by the condition that $\limsup P[M_p > c_n] < 1$ and $np^{-1}P[M_q > c_n] \to 0$ or the condition that $\limsup P[M_n > c_n] < 1$.

3. Proof of Theorem 2.1 and its corollaries. We begin with a lemma which is essentially the same in content and proof as results of Loynes (1965) and Leadbetter (1974).

LEMMA 3.1. *Let* (X_n) *have* AIM(c_n) *and let* (q_n), (α_n) *and* (p_n) *be as in Theorem* 2.1. *Let* $r_n = [n(p_n + q_n)^{-1}]$. *Then*

$$P[M_n \le c_n] \le P^r\left[M_p \le c_n\right] + o(1). \tag{3.1}$$

If, in addition, (2.3) *holds, then*

$$P[M_n \le c_n] - P^r\left[M_p \le c_n\right] \to 0. \tag{3.2}$$

We now present the proof of Corollary 2.2, which is the easier half of Theorem 2.1. We have

$$\begin{aligned} P[M_p \le c_n] &= 1 - P[M_p > c_n] \\ &= 1 - \sum_{i=1}^{p} P[X_i > c_n,\, M_{i,p} \le c_n] \\ &\le 1 - pP[X_1 > c_n,\, M_{1,p} \le c_n]. \end{aligned}$$

Applying (3.1), we obtain

$$\begin{aligned} P[M_n \le c_n] &\le P^r[M_p \le c_n] + o(1) \\ &\le \left(1 - pP[X_1 > c_n,\, M_{1,p} \le c_n]\right)^r + o(1) \qquad (3.3)\\ &\le \exp\{-nP[X_1 > c_n,\, M_{1,p} \le c_n]\} + o(1), \end{aligned}$$

since $prn^{-1} \to 1$.

The next step is the proof of the complementary inequality

$$P[M_n \le c_n] \ge \exp\{-nP[X_1 > c_n,\, M_{1,p} \le c_n]\} + o(1), \tag{3.4}$$

under the assumption that (2.3) holds. It is enough to prove (3.4) along subsequences such that $P[M_n \le c_n]$ converges, say to $L \in [0,1]$. By (2.3), $rP[M_p > c_n] \le C$ for some $C < \infty$ and for all n, so $P^r[M_p \le c_n] \ge (1 - Cr^{-1})^r \to e^{-C} > 0$. By Lemma 3.1, $L > 0$. If $L = 1$, (3.4) is obvious, so we can assume $0 < L < 1$. Let (s_n) be a sequence of positive integers such that $p_n = o(s_n)$ and $s_n = o(n)$; then (2.2) and (3.2) hold with p_n replaced by s_n and r_n replaced by $t_n := n(s_n + q_n)^{-1}$. By (3.2) in the two cases and the fact that $0 < L < 1$, $P[M_p > c_n] = o(P[M_s > c_n])$. Thus,

$$\begin{aligned} P[M_s > c_n] &= P[M_{s-p} > c_n,\, M_{s-p,s} \le c_n] + P[M_p > c_n] \\ &= P[M_{s-p} > c_n,\, M_{s-p,s} \le c_n](1 + o(1)) \\ &\le \left(\sum_{i=1}^{s-p} P[X_i > c_n,\, M_{i,p+i-1} \le c_n]\right)(1 + o(1)) \\ &\le sP[X_1 > c_n,\, M_{1,p} \le c_n](1 + o(1)). \end{aligned}$$

Since $nP[X_1 > c_n,\, M_{1,p} \le c_n]$ is bounded and $stn^{-1} \to 1$, we now deduce from Lemma 3.1 that

$$\begin{aligned} P[M_n \le c_n] &= P^t[M_s \le c_n] + o(1) \\ &\ge \left(1 - sP[X_1 > c_n,\, M_{1,p} \le c_n](1 + o(1))\right)^t + o(1) \\ &= \exp\{-nP[X_1 > c_n,\, M_{1,p} \le c_n]\} + o(1), \end{aligned}$$

as required.

We observe that the preceding arguments apply equally well along subsequences. The only remaining task is to show that the hypothesis (2.3) may be replaced by (2.4). It is enough to consider a subsequence along which $F^n(c_n) \to L_1$

and $P[M_n \le c_n] \to L_2$. If $L_1 > 0$, we are in the previous case. If $L_2 > 0$, then, by (3.3) $nP[X_1 > c_n, M_{1,p} \le c_n]$ is bounded. This fact and (2.4) together imply (2.3). Finally, if $L_1 = L_2 = 0$ and (2.4) holds, then (2.6) is immediate. □

4. Collections of sequences. It is often interesting to consider the simultaneous asymptotic nature of $P[M_n \le c_n]$ for (c_n) ranging over some collection of sequences. Frequently mentioned cases are the following one-parameter families of sequences: linear families $a_n x + b_n$, $x \in \mathbb{R}$, where $a_n > 0$, $b_n \in \mathbb{R}$ and the *quantile family* $(d_n(x))$, $x < 0$, given by $d_n(x) = \inf\{y\colon F(y) \ge 1 + n^{-1}x\}$, $-n < x < 0$ ($d_n(x) = -\infty$ for $x \le -n$). It was shown in O'Brien (1974b) and Leadbetter (1983) that if (X_n) has *extremal index* $\tau \in (0,1)$, that is, if $F^n(d_n(x)) \to e^x$ and $P[M_n \le d_n(x)] \to e^{\tau x}$ for all $x < 0$, then

$$P[M_n \le c_n] - G^n(c_n) \to 0 \tag{4.1}$$

for all sequences (c_n), where $G(x) = F^{\tau}(x)$. We will call any distribution function G satisfying (4.1) for all sequences (c_n) a *phantom distribution function* for (X_n). It is clear that G is not uniquely determined by (X_n) since only the right tail of G matters. We next show that (X_n) has a phantom distribution function even if the extremal index is zero or does not exist, including the case when there is no sequence c_n for which $F^n(c_n) \to a \in (0,1)$.

We first introduce a mixing condition which is a bit stronger than the AIM(c_n) condition if F is not continuous near x_F. The sequence (X_n) is said to satisfy AIM$^*(c_n)$ if it satisfies AIM(c_n) both as stated and with " $<$ " instead of " $\le$ " in (2.1).

THEOREM 4.1. *Let (X_n) have* AIM$^*(f_{[tn]})$ *for all $t > 0$, where $f_n := \inf\{x\colon P[M_n \le x] \ge e^{-1}\}$. Then (X_n) has a phantom distribution function G given by*

$$G(x) = e^{-1/n} \quad \text{if } f_n \le x < f_{n+1} \; (= 0 \text{ if } x < f_1). \tag{4.2}$$

PROOF. We note first that (f_n) is nondecreasing and $f_n \to x_F-$, so G is well defined. If $k = \max\{i\colon f_i = f_n\}$ and $m = \max\{i\colon f_i < f_n\}$, then

$$G(f_n) = G(f_k) = e^{-1/k} \ge e^{-1/n} > e^{-1/m} = G(f_m) = G(f_n-). \tag{4.3}$$

For $t > 0$, we have

$$P\big[M_{[nt]} \le f_{[nt]}\big] \ge e^{-1} \ge P\big[M_{[nt]} < f_{[nt]}\big].$$

Using an argument like that leading to (3.1) we can show from the first inequality that $np^{-1}P[M_q > f_{[nt]}] \to 0$, where (q_n) is a sequence arising from the AIM$^*(f_{[nt]})$ condition and (p_n) satisfies (2.2). Then, using this fact in lieu of (2.3), we obtain

$$P^t\big[M_n \le f_{[nt]}\big] + o(1) \ge e^{-1},$$

as in the proof of (3.2), so that

$$\liminf P\big[M_n \le f_{[nt]}\big] \ge e^{-1/t}. \tag{4.4}$$

Consider a subsequence along which $P[M_n < f_{[nt]}]$ converges. If the limit is positive, then $np^{-1}P[M_q \geq f_{[nt]}] \to 0$ as before and we obtain

$$\limsup P\left[M_n < f_{[nt]}\right] \leq e^{-1/t}, \tag{4.5}$$

along the subsequence and hence also for unrestricted n. Let (c_n) be any sequence of real numbers. It is enough to prove (4.1) for n restricted to subsequences along which

$$P[M_n \leq c_n] \to L \in [0,1]. \tag{4.6}$$

Assume first that $0 < L < 1$ and write $L = e^{-1/a}$ where $0 < a < \infty$. Let $\varepsilon \in (0,1)$. Comparing (4.6) with (4.4) for $t = a(1+\varepsilon)$, we see that $c_n < f_{[na(1+\varepsilon)]}$ for large n in the subsequence. Applying (4.3), we see that

$$G^n(c_n) \leq G^n\left(f_{[na(1+\varepsilon)]}-\right) \leq \exp\left\{-(a(1+\varepsilon))^{-1}\right\} + o(1).$$

Similarly,

$$G^n(c_n) \geq \exp\left\{-(a(1-\varepsilon))^{-1}\right\} + o(1).$$

Since ε is arbitrary, $G^n(c_n) \to e^{-1/a} = L$. The case $L = 0$ and $L = 1$ are handled similarly. □

We remark that the phantom distribution function G constructed in the above proof increases only by jumps even if (X_n) is i.i.d. and F is continuous.

We also remark that as an immediate corollary of Theorem 4.1, we obtain a new (albeit roundabout) proof of the fact that the set of possible limiting distribution functions of $a_n(M_n - b_n)$, $a_n > 0$, $b_n \in \mathbb{R}$, is the same for the stationary case (with suitable mixing) as it is in the independent case.

5. Functions of Markov sequences. We begin with a mixing condition that is appropriate for Markov sequences. Let $\sigma(i, j)$ denote the σ-field generated by $X_i, X_{i+1}, \dots, X_j$. If B is in this σ-field, say $B = \{(X_i, \dots, X_j) \in E\}$ for some Borel set $E \subset \mathbb{R}^{j-i+1}$, let B_k denote the shifted set $\{(X_{i+k}, \dots, X_{j+k}) \in E\} \in \sigma(i+k, j+k)$. We will say that (X_n) is *r-strongly mixing* (with mixing function g) for a positive integer r if

$$g(i) := \sup\left|r^{-1}\sum_{k=0}^{r-1} P(AB_k) - P(A)P(B)\right| \to 0 \quad \text{as } i \to \infty, \tag{5.1}$$

where the supremum is over all positive integers n, all $A \in \sigma(1, n)$ and all $B \in \sigma(n+i, \infty)$. Strong mixing as it is usually defined is the same as 1-strong mixing.

PROPOSITION 5.1. *Let (X_n) be stationary and r-strongly mixing and let (c_n) be a sequence of real numbers. Let (q_n) be any sequence of positive integers with $q_n \to \infty$ and $q_n = o(n)$. Then* (2.1) *holds, so (X_n) has* AIM(c_n).

PROOF. It is enough to consider subsequences along which $P[X_1 > c_n]$ converges. If the limit is 0, then

$$\begin{aligned}
&\left|P\left[M_i \le c_n,\, M_{i+q,\,i+q+j} \le c_n\right] - P\left[M_i \le c_n\right]P\left[M_j \le c_n\right]\right| \\
&\quad\le \left| r^{-1} \sum_{k=0}^{r-1} P\left[M_i \le c_n,\, M_{i+k+q,\,i+k+q+j} \le c_n\right] - P\left[M_i \le c_n\right]P\left[M_j \le c_n\right]\right| \\
&\qquad + rP[X_1 > c_n] \\
&\quad\le g(q) + rP[X_1 > c_n] \to 0.
\end{aligned}$$

If the limit is positive, then $P[M_i \le c_n] \le P[M_q \le c_n] \to 0$ which implies (2.1) in this case also. □

Obviously, r-strong mixing also implies the AIM* condition of Section 4.

We will deal with the theory of Markov sequences as described in the book by Revuz (1975). Let $(S, \mathscr{S})$ be a separable measurable space (often $S \subset \mathbb{R}^k$ for some k). Let $f\colon S \to \mathbb{R}$ be a measurable function and let (J_n) be a stationary positive Harris Markov (spHM) sequence with state space S. Let $X_n = f(J_n)$. The stationary sequence (X_n) is said to be a *function of the spHM sequence* (J_n). We recall that $P[J_n \in A$ infinitely often $|J_1] = 1$ a.s. for every set A for which $\pi(A) := P[J_1 \in A] > 0$. It follows in particular that (J_n) is irreducible. We will let T denote the transition kernel of (J_n) and let $\|\mu\|$ denote the total variation of μ.

If $S = \mathbb{R}$ and $f(x) = x$ then $(X_n) = (J_n)$ is itself an spHM sequence. Chernick (1982) described as open the question of whether every stationary Markov (X_n) satisfies $D(c_n)$ if $P[X_1 > c_n] \to 0$. Consider the following examples.

Let (Y_n) be an i.i.d. sequence with each Y_n uniform on $(0,1)$. Let W be a random variable which is independent of (Y_n) and satisfies $P[W = 0] = P[W = 1] = 0.5$. Let $c_n = 1 - n^{-1}$. If $X_n = \frac{1}{2}Y_n + \frac{1}{2}W$ for all n, then $D(c_n)$ fails, essentially because (X_n) is reducible. If $X_n = (-1)^{n+W}Y_n$ for all n, $D(c_n)$ again fails, this time because (X_n) is periodic. In the latter case (X_n) is 2-strongly mixing.

The following theorem is well known in the aperiodic case and is in fact an exercise in Revuz (1975). The extension to the periodic case is minor.

THEOREM 5.2. *Let (X_n) be a function of an spHM sequence with period r. Then (X_n) is r-strongly mixing with mixing function*

$$g(k) \le \int \left\| r^{-1} \sum_{i=0}^{r-1} T^{k+i}(x, \cdot) - \pi(\cdot) \right\| \pi(dx). \tag{5.2}$$

It follows from Theorem 5.2 that the various theorems on stationary sequences satisfying AIM(c_n) can be applied to the case of functions of spHM sequences. We now extend this observation to the case of functions of positive Harris Markov (pHM) sequences which are not necessarily stationary.

Theorem 5.3. *Let (K_n) be a pHM sequence on S and let $f\colon S \to \mathbb{R}$ be a measurable function. Let $Y_n = f(K_n)$ and let $R_{m,n} = \max(Y_{m+1}, \dots, Y_n)$ and $R_n = R_{0,n}$. Let (J_n) be the spHM sequence with the same transition probabilities as (K_n) and let $X_n = f(J_n)$ and $M_n = \max(X_1, \dots, X_n)$. Let (c_n) be a sequence of real numbers such that*

$$(5.3) \qquad P[Y_k > c_n] \to 0 \quad \text{for } k = 1, 2, \dots .$$

Then

$$(5.4) \qquad P[M_n \le c_n] - P[R_n \le c_n] \to 0.$$

Remarks. It follows that Theorem 2.1 holds for (Y_n) in the sense that if the hypotheses hold for the stationary sequence (X_n) rather for (Y_n) itself, then (2.5) holds with M_n replaced by R_n in the first term. It is not enough to assume $P[Y_1 > c_n] \to 0$ instead of (5.3) since this is compatible with $P[Y_2 > c_n] \not\to 0$. Note that (5.3) implies $P[X_1 > c_n] \to 0$ and, if the distribution of K_1 is absolutely continuous with respect to the distribution of J_1, the converse is also true.

Proof. Let $\mathscr{C}_1, \mathscr{C}_2, \dots, \mathscr{C}_r = \mathscr{C}_0$ be periodic classes of (J_n) so that $P[X_2 \in \mathscr{C}_{i+1} | X_1 \in \mathscr{C}_i] = 1$ for $i = 1, 2, \dots, r$. Let (L_n) be the pHM sequence with the same transition probabilities as (K_n) and (J_n), with $P[L_1 \in \mathscr{C}_i] = P[K_1 \in \mathscr{C}_i]$ for all i and with $P[L_1 \in A | L_1 \in \mathscr{C}_i] = P[J_1 \in A | J_1 \in \mathscr{C}_i]$ for all measurable sets A and all i for which $P[L_1 \in \mathscr{C}_i] > 0$. Let $Z_n = f(L_n)$, $T_{m,n} = \max(Z_{m+1}, Z_{m+2}, \dots, Z_n)$ and $T_n = T_{0,n}$. We will prove (5.4) by verifying

$$(5.5) \qquad P[R_n \le c_n] - P[T_n \le c_n] \to 0$$

and

$$(5.6) \qquad P[T_n \le c_n] - P[M_n \le c_n] \to 0.$$

To check (5.5), fix $\varepsilon > 0$ and choose m sufficiently large that

$$\| P[K_m \in \cdot\,] - P[L_m \in \cdot\,] \| < \varepsilon.$$

It follows from the Markov property that

$$\| P[(K_m, K_{m+1}, \dots) \in \cdot\,] - P[(L_m, L_{m+1}, \dots) \in \cdot\,] \| < \varepsilon$$

and hence that

$$\| P[(Y_m, Y_{m+1}, \dots) \in \cdot\,] - P[(Z_m, Z_{m+1}, \dots) \in \cdot\,] \| < \varepsilon.$$

Then, for n sufficiently large, we have

$$\begin{aligned} &|P[R_n \le c_n] - P[T_n \le c_n]| \\ &\quad \le |P[R_{m,n} \le c_n] - P[T_{m,n} \le c_n]| + P[R_m > c_n] + P[T_m > c_n] \\ &\quad < 2\varepsilon, \end{aligned}$$

where we have used (5.3) and the fact that the distribution of each Z_i is absolutely continuous with respect to the distribution of X_1. Since ε is arbitrary, (5.5) must hold. To prove (5.6), we note that for $i = 1, 2, \ldots, r$,

$$\begin{aligned}
&|P[M_n \le c_n | J_1 \in \mathscr{C}_i] - P[M_n \le c_n | J_1 \in \mathscr{C}_{i-1}]| \\
&\quad = |P[M_{1,n+1} \le c_n | J_1 \in \mathscr{C}_{i-1}] - P[M_n \le c_n | J_1 \in \mathscr{C}_{i-1}]| \\
&\quad \le P[X_1 > c_n | J_1 \in \mathscr{C}_{i-1}] + P[X_{n+1} > c_n | J_1 \in \mathscr{C}_{i-1}] \\
&\quad \to 0
\end{aligned}$$

by (5.3). It follows that

$$\begin{aligned}
&|P[T_n \le c_n] - P[M_n \le c_n]| \\
&\quad = \left| \sum_{i=1}^{r} \{P[T_n \le c_n | L_1 \in \mathscr{C}_i] P[L_1 \in \mathscr{C}_i] \right. \\
&\qquad \left. - P[M_n \le c_n | J_1 \in \mathscr{C}_i] P[J_1 \in \mathscr{C}_i] \right| \\
&\quad = \left| \sum_{i=1}^{r} P[M_n \le c_n | J_1 \in \mathscr{C}_i] \{P[L_1 \in \mathscr{C}_i] - P[J_1 \in \mathscr{C}_i]\} \right| \\
&\quad \le \sum_{i=1}^{r} |\{P[M_n \le c_n | J_1 \in \mathscr{C}_i] - P[M_n \le c_n | J_1 \in \mathscr{C}_1]\} \\
&\qquad\qquad \times \{P[L_1 \in \mathscr{C}_i] - P[J_1 \in \mathscr{C}_i]\}| \\
&\qquad + P[M_n \le c_n | J_1 \in \mathscr{C}_1] \left| \sum_{i=1}^{r} \{P[L_1 \in \mathscr{C}_i] - P[J_1 \in \mathscr{C}_i]\} \right| \\
&\quad \to 0. \qquad \square
\end{aligned}$$

We remark that Rootzén (1986) has studied functions of pHM sequences by means of regenerative cycles. In particular, he has obtained a different phantom distribution function for such sequences.

Many of the examples arising in the theory of extreme values of stationary sequences can be viewed as functions of spHM sequences. Some of these have the form $X_n = f(X_{n-1}, Y_n)$ for some function f, where (Y_n) is an i.i.d. sequence. For example, the autoregressive sequences in Chernick (1981) have this form and it is easy to verify that (X_n) is an aperiodic spHM sequence in those examples. The example of de Haan [cf. Leadbetter (1983)], which has properties similar to those of Example 5.3, satisfies $X_n = \max(X_{n-1} - 1, Y_n)$. We now consider three examples in detail.

EXAMPLE 5.1. This example shows the role of clustering of large values and shows that the relative sizes of the elements of the clusters play a role as important as the number of elements in the clusters. Let (Y_n) be an i.i.d. sequence with Y_1 uniform on $(-1, 0)$. Let $X_n = \max(Y_n, a^{-1}Y_{n+1})$ where $0 \le a \le 1$ ($a = 0$ corresponds to $X_n = Y_n$). Then (X_n) is a "moving function" of (Y_n) and is 1-dependent [strongly mixing with $g(2) = 0$]. Also (X_n) is a function of the spHM sequence $((Y_n, Y_{n+1}))$. If Y_n is large (i.e., close to 0), then X_{n-1} and X_n are both large, so large values come in pairs, although the values in the pairs are not equal for $a < 1$. It is easily checked that (X_n) has extremal index $(1 + a)^{-1}$. If g: $(-1, 0) \to \mathbb{R}$ is any strictly increasing continuous function, then $(g(X_n)) = \max(g(Y_n), g(a^{-1}Y_{n+1}))$ again has extremal index $(1 + a)^{-1}$, and g can be chosen so that the distribution function is in the domain of attraction of any prescribed extreme value distribution. The case $a = 1$ of this example was considered by Newell (1964).

Example 5.2. Let (X_n) be the spHM sequence with state space $S = [-1, 0)$ and transition probabilities as follows: given $X_1 \in S_1 := \{-1, -2^{-1}, -2^{-2}, \dots\}$, X_2 is uniform on S; given $X_1 = x \in S \cap S_1^c$, $X_2 = -2^{-k}$, where $-2^{-k} < x < -2^{-k-1}$. Thus (X_n) has period 2. There is no sequence (c_n) such that $F^n(c_n) \to a \in (0, 1)$. Nevertheless, $P[M_n \le xn^{-1}] \to \exp\{\frac{1}{2}x\}$ for $x < 0$ so (X_n) is in the domain of attraction of an extreme value distribution. Also, (X_n) has a phantom distribution function G given by $G(x) = 1 + \frac{1}{2}x$, $-2 < x < 0$.

Example 5.3. Let (X_n) be the stationary Markov chain with state space $S = \{1, 2, \dots\}$ and transition matrix T given by $T_{i,i+1} = t_i$ and $T_{i,1} = 1 - t_i$, where $0 < t_i \le 1$ for each i and $\mu := 1 + t_1 + t_1 t_2 + \cdots < \infty$. Let $H(x)$ be the distribution function of the first return time R to 1, starting at 1. Using the weak law of large numbers, one can deduce that $G := H^{\mu^{-1}}$ is a phantom distribution function for (X_n). If $t_i = t \in (0, 1)$ for all i, then $\mu = (1 - t)^{-1}$ and $G(x) = F^{(1-t)}(x)$. [Since $F^n(d_n(x))$ diverges, we cannot say (X_n) has extremal index $1 - t$.] Now suppose $t_i \to 1$ as $i \to \infty$. Then $P[M_{1,p} \le d_n(x) | X_1 > d_n(x)] \le P[X_2 \le d_n(x) | X_1 > d_n(x)] \to 0$, so (X_n) has extremal index 0, by Theorem 2.1. It is in general hard to specify the relationship between F and G beyond observing that $1 - G(x) = o(1 - F(x))$ as $x \to \infty$. If the function $1 - H(x)$ is regularly varying with exponent $\gamma < -1$ at ∞, then $x(1 - H(x))(1 - F(x))^{-1} \to -(\gamma + 1)\mu$. If, for example, $t_i = (i + 1)(i + 3)^{-1}$, then $1 - H(x) = 6([x] + 2)^{-1}([x] + 3)^{-1}$, $\gamma = -2$, $\mu = 3$ and $x(1 - H(x))(1 - F(x))^{-1} \to 3$.

Acknowledgments. I am grateful to E. V. Slud and J. Hüsler for some useful comments.

REFERENCES

Chernick, M. R. (1981). On strong mixing and Leadbetter's D condition. *J. Appl. Probab.* **18** 764–769.

Chernick, M. R. (1982). Letter to the editor. *J. Appl. Probab.* **19** 250.

Davis, R. A. (1979). Maxima and minima of stationary sequences. *Ann. Probab.* **7** 453–460.

Leadbetter, M. R. (1974). On extreme values in stationary sequences. *Z. Wahrsch. verw. Gebiete* **28** 289–303.

Leadbetter, M. R. (1983). Extremes and local dependence in stationary sequences. *Z. Wahrsch. verw. Gebiete* **65** 291–306.

Leadbetter, M. R., Lindgren, G. and Rootzén, H. (1983). *Extremes and Related Properties of Random Sequences and Processes.* Springer, Berlin.

LOYNES, R. M. (1965). Extreme values in uniformly mixing stationary stochastic processes. *Ann. Math. Statist.* **36** 993–999.

NEWELL, G. F. (1964). Asymptotic extremes for m-dependent random variables. *Ann. Math. Statist.* **35** 1322–1325.

O'BRIEN, G. L. (1974a). Limit theorems for the maximum term of a stationary process. *Ann. Probab.* **2** 540–545.

O'BRIEN, G. L. (1974b). The maximum term of uniformly mixing stationary processes. *Z. Wahrsch. verw. Gebiete* **30** 57–63.

REVUZ, D. (1975). *Markov Chains.* North-Holland, Amsterdam.

ROOTZÉN, H. (1986). Maxima and exceedances of stationary Markov chains. To appear in *Adv. in Appl. Probab.*

DEPARTMENT OF MATHEMATICS
YORK UNIVERSITY
NORTH YORK, ONTARIO M3J 1P3
CANADA